

Universal Multiple-Octet Coded Character Set
International Organization for Standardization
Organisation Internationale de Normalisation
Международная организация по стандартизации

Doc Type: Ideographic Research Group Document
Title: Response on Changing glyph and/or IDS of 97 KR Hanja chars containing U+53F1
Source: Eiso Chan (陈永聪, Culture and Art Publishing House)
Status: Individual Contribution to IRG #59, online meeting
Action: For consideration by IRG and ROK
Date: 2022-10-15

ROK submitted a response on [WS2021-00056](#) and other related issues. Thanks so much for Prof. Shin and Prof. Kim's hard works for so abundant examples and so detailed explanation. I don't want to argue for the linguistic issue at IRG meeting, but I need to clarify what I meant on the IRG ORT more clearly.

Firstly, we need to know how U+4E03/K0-7652 (七 칠), U+5315/K0-5D62 (匕 비) and U+2090E (匕 화) are written for ROK conventions.

Here is a *slanted horizontal bar running through the main hook* for U+4E03/K0-7652 (七). Please see Fig. 1.

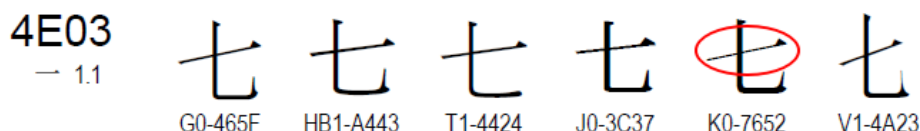


Fig. 1 U+4E03 on Code Charts

Here is a *slash not running through the main hook* for U+5315/K0-5D62 (匕).

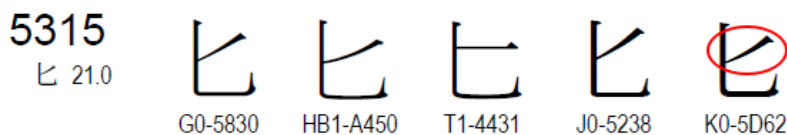


Fig. 2 U+5315 on Code Charts

There is no K-Source for U+2090E (匕) now, so I check the glyph picture on R. O. Korean CJK Ideographs Search System as below. Here is a *slash running through the main hook* for U+2090E (匕) like current G- and T-Source show. Please see Figs. 3 and 4.

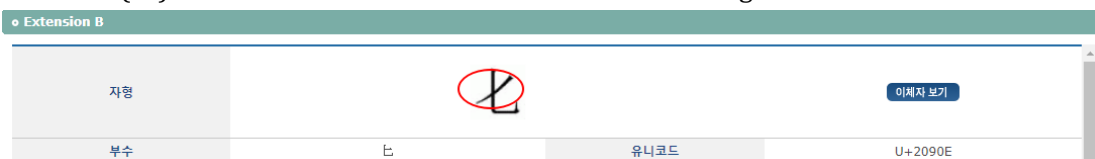


Fig. 3 U+2090E on CJK Ideographs Search System



Fig. 4 U+2090E on Code Charts

Secondly, we need know how U+53F1/K0-726A (叱 𪛗) is written for ROK conventions. As Fig. 5 shows, the right part of U+53F1/K0-726A (叱) is U+2090E (匕) like G-, H-, T-, J- and latent M-, KP-Sources not others. Only V-Source shows U+4E03 (匕) now. Please see Fig. 5. On the other hand, ROK shows the right part should be U+4E03/K0-7652 (匕) based on 《說文解字》, but there is no need to update the glyphs now.

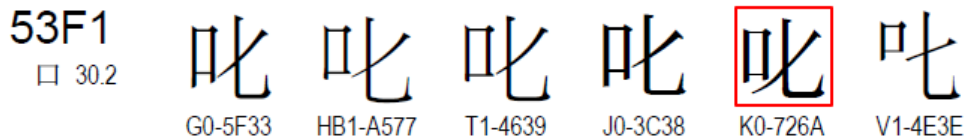


Fig. 5 U+53F1 on Code Charts

And then, the right part of U+20B9F (𪛗) is U+4E03 (匕), and the right part of U+2B738 (𪛗) is U+5315 (匕). Please see Figs. 6 and 7. Note that the first stroke of the right part of H-8D40 is a rectilinear horizontal bar, that is different from HB1-A443 showed in Fig. 1, but it is not doubtful to identify the right part as U+4E03 (匕); the second first stroke of U+5315 (匕) is the common horizontal bar for TCA (even latent Macao SAR) conventions.

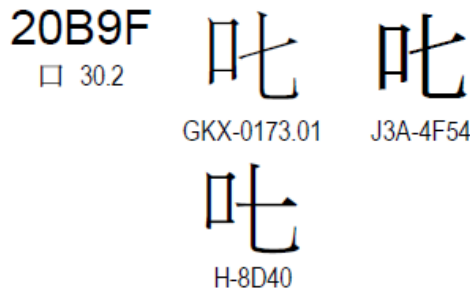


Fig. 6 U+20B9F on Code Charts



Fig. 7 U+2B738 on Code Charts































































Thirdly, we need to know the compound rule for this kind of compound Hanjas for discussion, which are only used to record Korean language and the dialects in Korean Peninsula, Chinese Chosen People and Central Asian Korean People.

As ROK's response shows, there are two types of compound modes, at least for this issue. The first type includes the common Hanja as the top to indicate the initial consonant (choseong) and the medial vowel(s) (jungseong), and the component as the bottom to indicate the tail consonant (jongseong). The reading value of the top is based on the original Korean reading, and the one of the bottom is to use the systematic rule. (See 3.1 and 3.2 in ROK's response.) The second type includes the component as the top to indicate the fortis sound mark for the initial consonant, and the common Hanja to indicate a complete unchanged syllable. The reading value of the top is to use the systematic rule, and the one of the bottom

is based on the original Korean reading. (See 3.3.) Commonly, there is only one systematic component for the second type, which is the focus of this discussion.

To sum up, there is only one systematic component used as tail consonant or the fortis sound mark in one character.

Fourthly, we need to confirm the rationale of WS2021-00056:KC-00023 as below.

00056	5.0 乙	 KC-00023																																																																																																																																																																																																																																																									
-------	-------	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	--	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	--

301D4 □ 30.6	𪛗 KC-00414				
301D9 □ 30.6	𪛘 KC-00411	30205 □ 30.9	𪛙 KC-00458	30217 □ 30.10	𪛚 KC-00493
301E4 □ 30.7	𪛛 KC-00434	3020C □ 30.9	𪛜 KC-00474	30220 □ 30.11	𪛝 KC-00502
301E8 □ 30.7	𪛞 KC-00436	30211 □ 30.10	𪛟 KC-00486	30234 □ 30.13	𪛠 KC-00523
301FF □ 30.9	𪛢 KC-00472	30215 □ 30.10	𪛣 KC-00480	30235 □ 30.13	𪛤 KC-00547
30200 □ 30.9	𪛦 KC-00476	30216 □ 30.10	𪛧 KC-00490	30237 □ 30.14	𪛨 KC-00546
517A 八 12.8	𪛩 GK-683F			𪛩 K2-2426	
5391 厂 27.5	𪛪 GK-6876			𪛪 K2-2576	
54DB □ 30.6	𪛫 GK-685E			𪛫 K2-2729	
551C □ 30.7	𪛬 GE-3356			𪛬 K0-5845	
551F □ 30.7	𪛭 GK-677E			𪛭 K2-2747	

Fig. 10 Screenshot of the examples in ROK's response

According to Fig. 10, we can know the U+5315/K0-5D62 (𪛬) forms are the most common, but

some characters also use the U+2090E (匕) form, such as U+2D1CD (𠩺), U+2D1DD (𠩻) and U+2D2AC (𠩼) in Ext. F.



Fig. 11 Screenshot of the examples in ROK's response

Now, we can find the glyphs of this component are not consistent, and which one is more common. If ROK hopes to keep the IDS element as U+53F1/K0-726A (𠩺), it is better to update this components of all the examples mentioned in ROK's response as U+2D1CD (𠩺), U+2D1DD (𠩻) and U+2D2AC (𠩼); if not, it is better to update this component of U+2D1CD (𠩺), U+2D1DD (𠩻) and U+2D2AC (𠩼) to the common one later, and change the IDS element to U+2B738 (𠩺). I think the latter one will be more sensible. China and Japan also use the U+2B738 (𠩺) form for these kind of compound ideographs, such as U+2E0BB (𠩺) and U+2E1CE (𠩺) for J-Source uses.

Also note that IDS is not to show the real philological rationales in IRG encoding works now, but the intuitive descriptions of the compound information of the components and the positions in one character. This is the reason why we don't use 𠩺 禾 as the intuitive IDS of 穎 but 𠩺 禾 頁.

Finally, if ROK experts are interested in the IDSes to show the real philological rationales in future, the new IDCs submitted in [IRGN2572](#) will be useful and needed. Some components are omitted when the whole character is used as the element of other more compound character. Based on 《說文解字》, the right part of U+9689/K2-6B79 (隍) is the omitted form of U+6BC0/K2-3E73 (毀), so the philological IDS should be 𠩺 𠩺 & SST; 毀 𠩺 (SST means IDEOGRAPHIC DESCRIPTION CHARACTER COMPONENT SUBTRACTION here, see [my feedback](#) on [IRGN2492](#)).

(End of Document)